

APPENDIX D

DIGITAL WAVE DATA COLLECTION AND ANALYSIS PARAMETERS

D-1. Purpose. This appendix describes the parameters used in the collection and analysis of digital wave data. The selection of appropriate sampling and analysis parameters is essential for a successful data collection program. Detailed information is given in item 132.

D-2. Duration. Duration is the total time data is collected. Duration is typically measured in days, months, or years. Several years of data are necessary to discern annual trends or to make extremal predictions.

D-3. Burst Interval. Burst interval is the time between sample records. Sample records may be recorded continuously or every few hours (typically every 1, 2, 3, 4, or 6 hours). Also, a threshold may be defined so that data are collected continuously during storm conditions but only intermittently during calm conditions.

D-4. Sampling Frequency, Sampling Interval, and Nyquist Frequency. These three sampling parameters are interrelated, so choosing one of the three determines the other two.

a. The sampling frequency f_s (in Hertz) is related to the sampling interval Δt (in seconds) by

$$f_s = \frac{1}{\Delta t}$$

One Hertz is one sample per second. Typical values of the sampling frequency are 1, 2, or 4 Hz. The Nyquist frequency is the highest frequency that can be detected when sampling at the selected sampling frequency. The Nyquist frequency f_{ny} is defined

$$f_{ny} = \frac{1}{2\Delta t} = \frac{f_s}{2}$$

b. Two undesirable phenomena, aliasing and hidden oscillations, can occur when sampling at a constant rate. Aliasing is the folding back of energy from frequencies higher than the Nyquist frequency into frequencies related to harmonics of the Nyquist frequency, i.e.,

$$(2 f_{ny} + f), (4 f_{ny} + f), \dots (2n f_{ny} + f) \quad n = 1, 2, 3 \dots$$

where f is any frequency between zero and the Nyquist frequency. Hidden oscillations are the loss of kinetic energy at a particular frequency because the same point in the cycle of the process is always sampled; therefore, the information about the cyclic nature of the process is lost.

c. Three methods to prevent these undesirable phenomena are:

7 Jul 89

(1) Reduce the higher frequencies present to less than the Nyquist frequency by low-pass filtering the signal prior to digitization with an analog, anti-aliasing filter.

(2) Randomly vary the sampling interval such that the sampling interval approaches a uniform distribution.

(3) Select a constant sampling interval at least twice the highest frequency component present.

The third method is used most often because it does not require special equipment. Generally, high frequencies contain relatively little energy, so they are of little interest. Typically the upper frequency limits of interest are 0.35 Hz for ocean waves, 0.50 Hz for waves in bays and lakes, and 0.25 Hz for low-frequency harbor oscillations.

D-5. Total Number of Points, Record Length, and Frequency Increment.

a. The total number of data points N and the record length T_r (in seconds) are related by

$$T_r = N\Delta t$$

Traditionally, the total number of data points has been a power of 2 (typical values are 1,024, 2,048, and 4,096) because fast fourier transform (FFT) routines to transfer the data from the time domain time series to the frequency domain wave spectra required it. Now FFT's are available in multiples of powers of 2, 3, and 5. Typical values of the record length are 17 to 68 minutes. Longer record lengths give higher resolution and greater confidence in the spectral estimates, but the environment conditions must not change significantly during the sample.

b. The frequency increment Δf (in Hertz) in the frequency domain is analogous to the time domain sampling interval Δt

$$\Delta f = \frac{1}{T_r}$$

Wave energy density spectra are calculated from the measured time series at discrete values which are integer multiples of the frequency increment. There will be $N/2$ wave energy density values ranging from Δf to the Nyquist frequency, where N is the total number of data points in the time domain.

D-6. Number of Averages, Resolution Bandwidth, and Degrees of Freedom. These last three parameters are concerned with analysis of the data after it is collected.

a. Energy density values are estimates of the true wave spectrum. An infinite number of data points and an infinite number of samples would be required to calculate the true energy density. Since this is impossible, the spectral estimates are usually averaged in the time domain (ensemble averaging) or the frequency domain (band averaging) to increase the confidence in

the estimate; but as confidence is increased by averaging, resolution is lost. If the raw spectral estimates are band averaged in the frequency domain, the number of average numbands is related to the resolution bandwidth B (in Hertz) by

$$B = \frac{\text{numbands}}{T_r}$$

Typical values of the number of averages are 8 and 16. The corresponding resolution bandwidth is 0.00781 Hz and 0.01563 Hz for a 1,024-second record length.

b. If spectral estimates are assumed constant within the bandwidth, they are considered chi-square variables with degrees of freedom nu given by

$$\text{nu} = 2 \text{ numbands}$$

The number of degrees of freedom is used to calculate the confidence intervals on the autospectral energy density estimates. The larger the nu value, the tighter the confidence intervals for a given record length. Typical values for the number of degrees of freedom are 16 and 32 for bandwidths of 8 and 16, respectively.